Author's response to reviews

Title: Mental health problems in the 10th grade and non-completion of upper secondary school: the mediating role of grades in a population-based longitudinal study

Authors:

Åse Sagatun (aase.sagatun@r-bup.no)
Sonja Heyerdahl (sonja.heyerdahl@r-bup.no)
Tore Wentzel-Larsen (tore.wentzel-larsen@r-bup.no)
Lars Lien (lars.lien@medisin.uio.no)

Version: 4
Date: 22 November 2013

Author's response to reviews: see over
Dear Editor

Thank you for valuable comments to our manuscript “Mental health problems in the 10th grade and non-completion of upper secondary school: the mediating role of grades in a population-based longitudinal study”.

Below is a point-by-point response to the concerns.

Best regards, Åse Sagatun

Reviewer: Theis Lange

General comments:

The paper is generally well written with clear objectives, results, and conclusions. The analysis combines a large high quality dataset with suitable statistical tools, including fairly recent tools for causal mediation analysis. 

Major Compulsory Revisions

1) The conclusion as it is written in the Abstract promises more than the paper can deliver. The potential serious problem of reverse causation (which is discussed in the last paragraph of the subsection “Methodological Strengths and Limitations” in the Discussion section of the main text of the paper) should be acknowledged.

Authors’ response: The conclusion in the abstract is rewritten. We now state clearly that our assumption is that the direction of the effect is from mental health problems at age 15-16 to school dropout in 10th grade. We did not want to have a discussion as part of the conclusion.

“Assuming a causal relationship from mental health problems to school performance, this study shows that externalising problems impair educational attainment”

2) The causal mediation analysis is very depended on the causal connections between the two measures of mental health (ie. internalizing and externalizing). In the reported mediation analysis the two measures are controlled for each other; so that in the analysis with internalizing as the exposure the degree of externalizing is treated as a covariate and vice versa, however, no arguments for this choice are given. If in reality one of the measures causes the other controlling for the other measure might bias the mediation analysis. Unless the two measures of mental health problems are independent conditional on baseline covariates there is no easy solution to this problem, but it should at the very least be acknowledged and discussed.

Authors’ response: We have now done the mediation analysis for externalising problems without adjusting for internalising symptoms and vice versa (Model 1 in Table 5). The models in the Logistic regression and Mediation analysis are now the same. The rationale behind the models is explained in the Methods
section, page 10 last paragraph: “First we studied the effect of internalising and externalising problems separately (Model 1). In order to study the independent effect of externalising problems we adjusted for internalising scores and vice versa (Model 2). In the final model we also included health behaviours (Model 3).”

We have also performed partial correlations between internalizing and externalizing problems, controlling for health behaviours and sociodemographic variables. The partial correlation among boys was 0.307, among girls 0.375. These results are very similar to the Pearson correlations which were 0.265 and 0.350 respectively. We have reported the correlations between the mental health scales, health behaviours and mental health and health behaviour collectively in the paper (page 13).

The uncertainty about a causal connection between the two mental health measures is acknowledged in the “Strengths and limitations” part of the discussion: “We found a weak to moderate correlation between externalising and internalising problems but a causal relationship cannot be excluded. However, we have studied both the two variables separately and controlled for each other. In the causal mediation analyses, the estimates for externalising problems were quite similar in the models without versus with adjustment for internalising problems, while the estimates related to the internalising scale became weaker when adjusting for externalising problems (Model 1 and 2). We cannot rule out that we might have underestimated the effect of internalising symptoms by controlling for externalising problems.”

3) The Tables 4 and 5 both consider the effect of internalizing and externalizing on the probability of non-completion. However, in Table 4 the effect of an increase of 10% of range is presented while in Table 5 the effect corresponding to a 2 point change is presented. This, as far as I can tell unnecessary difference, makes the tables hard to compare. The authors should consider sticking to one definition of change in internalizing and externalizing scores.

Authors’ response: The ranges of the SDQ scores are 0-20, and a 10% change corresponds with a 2 points change. We acknowledge the unnecessary difference in the presentation and have changed the labels in Tables 4 and 5.

4) In the last part of the subsection “Statistics” in the methods section a sensitivity analysis of the causal mediation analysis is discussed, including a correlation parameter #. I was unable to understand what was done in this sensitivity analysis. This also applies to the reporting of the sensitivity analysis in second last paragraph of the result section. In both places extended explanation is needed.

Authors’ response: The description of sensitivity analyses in the Statistics section (page 10) has been expanded considerably, and the reference [27] has been replaced by an article giving more details on the role of sensitivity analysis in causal mediation analysis.

“For model identification, the causal mediation analysis makes the untestable sequential ignorability assumption, that in particular implies that there are no unmeasured confounders [27]. Briefly, sequential ignorability means that conditional on the confounders, actual treatment (corresponding
to internalising, resp. externalising problems in our case) is independent on all potential values of outcome and mediator, and, also conditioning on actual treatment, the actual mediator is independent on all potential values of the mediator. Potential here refers to a counterfactual setting. Since this assumption is untestable sensitivity analysis is recommended. Specifically [27, Theorem 2 page 316], model identification is also possible under deviations from sequential ignorability, assuming a specific correlation $\rho$ between the error terms in the models for the mediator and the outcome, with zero correlation corresponding to sequential ignorability. Such an error correlation is reasonable when there are unmeasured confounders for outcome and mediator. When the direct and indirect effects have the same sign for a large $\rho$ interval, the sensitivity of the conclusions of the causal mediation analysis is considered as small. “Old reference 27: Imai K, Keele L, Tingley D, Yamamoto T: Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. Am Polit Sci Rev 2011, 105:765-789.


Minor Essential Revisions

5) In background section: Change “Thus, adjusting for these predictors in important when causal models are tested” to “Thus, adjusting for these predictors in important when causal models are employed”.

Authors’ response: In accordance whit the other reviewer’s point 3, the sentence referred to is removed from the background section.

6) The first sentence of the subsection “The present study” discusses internalizing and externalizing, but these concepts have yet not been defined.

Authors’ response: We have changed the first sentence of the subsection “The present study” to: “We study the extent to which mental health problems, both internalising (emotional and peer problems) and externalising problems (conduct problems, hyperactivity-inattention), in the 10th grade (age 15–16 years) are associated with non-completion of upper secondary school during the following 5 years.”

7) In the second sentence of the method section is discussed that health surveys were conducted between 1999 and 2004, but it is not stated how many waves were conducted.

Authors’ response: The health surveys were conducted in the 10th grade in six Norwegian counties in the time period 1999-2004. This may have been unclear in the text, and the sentence is now changed. We have also removed the first cohort from the description of the study, because it is not included in our analysis. The second paragraph (page 5) is now changed: “The Norwegian Institute of Public Health conducted these comprehensive health surveys in regional cohorts between 2000 and 2004.”
8) In the subsection “Measures” please explain what the “final 11 grades” are.

Authors’ response: We acknowledge that the sentence may have been unclear and it is now changed to: “The final grades (0–6) in the 11 main school subjects from 10th grade are summed to the variable “grade points” (0–66).”

9) In the subsection on “Mental health problems” it would be great if the authors could provide some intuition for the two subscales (internalizing and externalizing).

Authors’ response:

We have added a few words about the internalizing and externalizing scales, to provide some information in relation to the content of the scales (page 7): “There is theoretical and empirical support for combining the SDQ’s emotional symptoms and peer problems subscales into an ‘internalising’ subscale, covering anxiety, depression and withdrawal. The conduct problems and hyperactivity-inattention subscales are combined into an ‘externalising’ subscale. In the current paper, we used the internalising and the externalising subscales with a possible range of scores from 0 to 20 with higher scores representing more problems [20].” We have also added more information on the scoring procedure (in line with the other reviewer’ point 7):

“Each item can be answered with “not true” (0), “somewhat true” (1) or “certainly true” (2), with reference to the past 6 months. For each subscale these values were summed to generate scale scores ranging from 0 to 10 [19].”

10) In the subsection “Statistics” the concepts direct and indirect effects are used without any introduction or explanation. I would suggest adding a few lines on this.

Authors’ response: Where we introduce the causal mediation analysis we have now added (page 10): “Indirect effects are the effects mediated via grades, and the direct effects are the effects not mediated.”

11) In the legend to “Additional File 1” please change “tested” to “assumed”.

Authors’ response: We have changed “tested” to “assumed”.

Discretionary Revisions

12) The Tables 4 and 5 both consider the effect of internalizing and externalizing on the probability of non-completion, but Table 5 reports absolute effects while Table 4 reports relative effects (ORs). This makes comparison unnecessarily difficult. I would suggest sticking to a single type of effect measure (e.g. OR). If the choice in Table 5 is dictated by the capabilities of the employed mediation-package

**Authors’ response:** We agree that it would have made comparisons between table 4 and 5 easier if we had used the same type of effect measures in the two tables, but the software we have employed (mediation package in R) does not give us the relative effect. However we think that the analysis performed are well suited to answer our research questions, and we have not done any new comprehensive analysis.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.
Reviewer: Saskia van Dorsselaer

General comments:
Well written paper with clear use of language. Good Summary of the results. I recommend to accept this paper for publication with minor revision.

#1 Background
Discretionary Revision: The authors state that the association between mental health problems and educational outcome through different pathways is not well studied. No references are given. Is it not studied at all? In the discussion there are some references on this topic (under 'Mental health and subsequent educational achievement'). Please integrate these references more in the Background-part of the article.

Authors’ response: When we in the background state that the association between mental health problems and educational outcome through different pathways is not well studied, we are thinking of studies that elaborate on possible mechanisms for this association i.e. by mediating analysis. In the discussion part we are referring to studies that are looking at associations, without using mediating analysis. We acknowledge that this difference is unclear, and have now made this clearer in the introduction by replacing the sentence “This is not well studied” with “Studies elaborating on possible mechanisms for the association between various mental health problems and educational outcome are hard to find.”

#2 Minor Revision: (Last paragraph) 'The risk literature for externalising and internalising problems and academic attainment provides clues about potential common causes, such as sociodemographic factors and health behaviour [13, 15, 16].' Please tell more about the content of these references, e.g. which health behaviour is related to mental health problems and how they relate to the health behaviour measured in this paper.

Authors’ response: We have included examples on what kind of health behavior that is measured in the papers referred: “The risk literature for externalising and internalising problems and academic attainment provides clues about potential common causes, such as sociodemographic factors and health behaviour like physical activity, smoking and alcohol habits [13, 15, 16].”

#3 Discretionary Revision: (Last sentence) 'Thus, adjusting for these predictors is important when causal models are tested.' This sentence is more appropriate for the method section.

Authors’ response: The sentence is now removed.
#4 Methods

Minor Revision (first and second paragraph): In this paper only the four studies between 2000 and 2004 were used. Why mention the one in 1999 and include it in the overall N of 15,966? Now the authors give information about a part of the sample that was never used in the first place. It would be more straightforward to count only the sample which was used and report the response for that.

**Authors’ response:** We acknowledge that a description of the whole study is unnecessary and have rewritten the method section to only describe the study population used in the present paper.

“The overall response rate of the questionnaire studies was 87% (n = 12,434). Of the participants in the baseline studies, 88% (n = 10,931) accepted linkage of information between the survey and official registers.”

#5 Minor/major Revision (second paragraph): the authors say on several places (e.g. in the discussion) in the manuscript that little is known about the 12% that did not accept linkage of information. However without the linkage it is easy to analyze the differences between students who did and those who did not accept the linkage of information on age, gender, SDQ-scores, grade point at grade 10 and give a short report on these differences. Although it is only a small proportion who did not accept the linkage, I think this should be part of the description of the sample. Therefore I would like to see such a analysis described in text.

**Authors’ response:** We have now compared mean values of the SDQ scales and gender distribution among those who accepted linkage of information and those who did not, and no statistically significant differences were found between the two groups. In the description of the data we have now included these analyses. At the end of the paragraph describing the “Study population and procedure” we have now added: “The 12 percent that did not accept linkage to the national registers did not differ significantly in gender distribution as our study sample (p = 0.782). The two groups differed neither significantly in symptoms of internalising (p = 0.581) nor externalising problems (p = 0.164)”.

The participants had the same age since all the studies were performed in the 10th grade. Information about grades used in our analysis was from Statistics Norway and only available for those accepting linkage and therefore not possible to compare between the two groups.

#6 Measures, The Norwegian School System (very welcome information for the understanding of the paper and method): Minor Revision (first paragraph): I think the sentence 'As early as 1920, the Principles of a common school for all were adopted [17].', is too detailed and could be leftout.

**Authors’ response:** The sentence 'As early as 1920, the Principles of a common school for all were adopted [17].' is removed from the paper.
#7 Measures, Mental health problems

Minor Revision (last paragraph): 'On a 0–20 scale, this corresponds to a difference of 2, which is the difference we use in all regression and mediation analyses." I understand from this sentence that the SDQ-scores are not calculated on a 0 to 20 scale but with steps of 2, so 0-2-4-6 etc. Is this correct?

In different parts of the document the data of the SDQ is described differently. In the part on 'Statistics' I read: 'Direct and indirect effects were studied for a 2-point increase in SDQ score, from 4 to 6 for internalising problems and from 6 to 8 for externalising problems; these intervals were chosen within the main range of the respective distributions. In the tables it is as "per 10% of range" This is very unclear to me. Please write more clearly how SDQ-scores are used in analysis.

Authors' response: The SDQ scales (internalising and externalising problems) range from 0-20 with one-point intervals. More details on the scoring is now included in the presentation of the SDQ: “Each item can be answered with “not true” (0), “somewhat true” (1) or “certainly true” (2), with reference to the past 6 months. For each subscale these values were summed to generate scale scores ranging from 0 to 10 [19]. There is theoretical and empirical support for combining the SDQ’s emotional symptoms and peer problems subscales into an ‘internalising’ subscale, covering anxiety, depression and withdrawal. The conduct problems and hyperactivity-inattention subscales are combined into an ‘externalising’ subscale. In the current paper, we used the internalising and the externalising subscales with a possible range of scores from 0 to 20 with higher scores representing more problems [20].”

Because a difference of about 10 points on a 0–100 scale is often considered noticeable [21] we have used a 2-point interval in the presentation of the regression and the mediation analyses. On a 0–20 scale, 2 points correspond with 10 percent. In table 4 “per 10% of range” is changed to “per 2 points”.

#8 Measures, Sociodemographic factors and family background

Minor Revision (second paragraph, last sentence): 'Members of the Sami population, the indigenous population from northern Norway, are thus defined as ethnic Norwegians if at least one of the parents is born in Norway.' This sentence is abundant since in the first sentence of this paragraph the definition of ethnic minorities is given as 'those having both parents born in a country other than Norway.

Authors' response: The sentence about the Sami population is now removed.

#9 Statistics

Minor issues not for publication: Word missing in the sentence "To determine whether the patterns between mental health and non-completion of upper secondary school were the same in both general and vocational tracks, we re-ran the logistic regressions and causal mediation analyses by stratification on type of track.
Authors’ response: We have rewritten the end of the sentence: “To determine whether the patterns between mental health and non-completion of upper secondary were the same in both general and vocational tracks, we re-ran the logistic regressions and causal mediation analyses for the two tracks separately.”

#10 Results, background characteristics

Minor Revision (first paragraph): Please add the differences between vocational and general tracks in Table 1.

Authors’ response: General and vocational tracks are added in Table 1, and the figures in the text are removed and instead referred to in Table 1.

“The percentage starting at vocational tracks was 52.8 among boys and 41.0 among girls. For both genders, the prevalence of not completing upper secondary school was about three times higher for vocational tracks (Table 1).”

#11 Minor Revision (second paragraph): Please use small p for p-values and not capitals. (unless this is journal policy)

Authors’ response: Capital P is changed to small p throughout the paper.

#12 Major revision (second paragraph): For a good mediation analysis it is necessary to establish the association between all components of the mediation. I miss the association between SDQ and grades. Maybe an extra table is needed with this association and the one between grades and completion as described in this paragraph. I would be satisfied with a description in text.

Authors’ response: In the end of the section “Background characteristics” the final grades (in 10th grade) are described by completion of upper secondary school and gender. In the end of the paragraph “Background characteristics” we have now added: “Mental health problems were also negatively associated with grades [(Pearson correlation, r): SDQ externalizing, boys (-0.38, p<0.001), girls (-0.41, p<0.001) and SDQ internalising, boys (-0.13, p<0.001), girls (-0.24, p<0.001)].”

#13 Results, The association between psychiatric problems in the 10th grade, Health behaviour in the 10th grade and non-completion of upper secondary school

Authors’ response: This seems to be a repetition of point #14. This is a heading in the manuscript, and there is no comment to this
#14 Minor Revision: "The association between psychiatric problems in the 10th grade, health behaviour in the 10th grade and non-completion of upper secondary school".

Please do not use the term psychiatric problems but mental health problems here. This is measured with the SDQ-selfreport and although the authors discuss the linkage between SDQ and psychiatric problems, I think the term is too strong here.

Authors' response: The term “psychiatric problems” is changed to the term “mental health problems” throughout the paper.

#15 Minor revision: First paragraph: A lot of information not in the table. If you turn table 2 the information on the difference between vocational and general track fits in the table.

Authors' response: We have now stratified table 2 by vocational and general track, and removed the figures in the text.

#16 Minor revision: Third paragraph "The ORs for externalising problems in this model were reduced in both genders [boys; 1.27 (1.21–1.33) and girls; 1.31 (1.24–1.39)]." Please add: but still significant.

Authors' response: “but still significant” is added in the end of the third paragraph in the section "The association between mental health problems in the 10th grade, health behaviour in the 10th grade and non-completion of upper secondary school".

#17 Minor revision: last paragraph:

"In the final model (Model 3), smoking, alcohol use and physical activity remained independent predictors for non-completion of upper secondary school (Table 4a and b, Model 3)." Alcohol use and physical activity were only significant for boys. Please make this difference.

Authors' response: We have changed the last paragraph to: “In the final model (Model 3), smoking, alcohol use and physical activity remained independent predictors for non-completion of upper secondary school among the boys (Table 4a, Model 3). In girls, smoking and physical activity remained independent predictors (Table 4b, Model 3).”

Results, The mediating role of grade points when analyzing etc.

#18 Minor revision (last paragraph): "However, the sensitivity analysis for direct effects of externalising problems showed low robustness (#< 0.2 for girls and #< 0.1 for boys)." It is not discussed what the consequences of these findings are for the interpretation of the results. Please give more information, here or in the discussion.

Authors' response: See our answer above to the other reviewer (point 4) on sensitivity analysis, included more details in the Statistics section on the interpretation of the results of sensitivity
analysis. We have also expanded the description of the results of the sensitivity analysis in Discussion (sentence 4 from the bottom, page 18): “In sensitivity analyses, our results seem to be robust for moderate residual correlations, corresponding to a moderate amount of unmeasured confounders in the models for the mediator and outcome.”

**Discussion, methodological strengths and limitations**

**#19** Minor revision (Second paragraph): “Even though we might have a selective loss to follow-up”. If you do additional analysis as mentioned before, more can be said about this, here.

**Authors’ response:** We have discussed the possible problem with selection bias more based on the new analysis described in the method.

In the section “Methodological strengths and limitations” we have now written: “A limitation of the study is that not all those invited took part in the baseline study, and not all participants accepted the survey to be linked with national registers. Thus our analyses are based on 77% of all 10th graders in the respective counties for the years in question. Of the participants in the baseline study 12 percent did not accept linkage of their survey. Even though the participants not accepting linkage did not differ in their report of symptoms of mental health problems we might have a selective loss to follow-up. However the percentage in our study not completing upper secondary school within 5 years is in accordance with Statistics Norway’s national numbers for the comparable years (68%–72%)[17]. There is also some empirical support that generalisation for measures of association is less sensitive to loss to follow-up than prevalence measures [28]. A paper based on parts of the same sample considers response rates and selection problems by investigating mental health and health behaviour variables [28]. Here the association measures (Prevalence ratios) were quite similar among participants and all invitees. The response rate was quite high, and we expect that the findings are fairly representative for the study population.”

**#20** Minor revision (Second paragraph): “Nevertheless, early conduct problems have been found to be associated with lower IQ and attention difficulties, with the result that the difficulties faced by children with conduct problems were exacerbated by lower average IQ and higher rates of attentional problems [31].” If this is true, more differences would been found between the role of externalizing problems and school drop-out between the different educational tracks. This is not the case. Please reflect on this.

**Authors’ response:** We do not quite understand this comment. We found that the levels of symptoms of mental health problems and school dropout differed between the two tracks, but that the associations were similar.

**Discussion, Our results according to previous findings etc**
Minor issues not for publication: "Externalising problems (conduct problems and hyperactivity–inattention) may reflect a cumulative effect of inattention and learning across the schooling career, and may have an impact on the acquisition of academic skills. Conduct disorders lead to repeated disciplinary actions, which are likely to affect students’ engagement with schooling and to influence their grades. Internalising problems (symptoms of depression, anxiety and peer problems) are likely to disrupt students’ overall social functioning and perceived competence, leading to diminished motivation, but may be less associated with academic skills."

Another possibility is that externalizing problems have a more stable nature, they tend to stay over a longer period of time during adolescence and even through adulthood, whereas internatilising problems may have of shorter term nature and may disappear when an individual changes school (peer problems).

Authors’ response: We appreciate this comment, which adds to our reflections on the topic. We have revised the paragraph, including the reviewer’s point.

"Externalising problems (conduct problems and hyperactivity–inattention) have a more stable nature and may reflect a cumulative effect of inattention and learning across the schooling career, which may have an impact on the acquisition of academic skills. Conduct disorders lead to repeated disciplinary actions, which are likely to affect students’ engagement with schooling and to influence their grades. Internalising problems (symptoms of depression, anxiety and peer problems) are likely to disrupt students’ overall social functioning and perceived competence, leading to diminished motivation, but may have a shorter term nature, and may be less associated with academic skills."

#22

Table 1

Minor issues not for publication: Please remove the text "sample drop out rate (32,4)" It is confusing since it regards the total sample and not boys and girls and this exact information is also written in the text.

After ‘Father’s income at age 16’ and ‘Mother’s income at age 16’ there is an * but this is not explained. Explain or remove the asterix.

Authors’ response: We have now removed the sample drop out rate (32.4) and * from the table 1.

Table 2 to 4

Minor revision: Please mark significant differences between Completers and non-completers in the table, tests (t-test, Spearman Rank test, logositc regression analysis) and level of significance in subscript. A reader should be able to read a table without much of the additional text.

Table 5 P < 0.005 should be p< 0.05?
Authors’ response: We have now marked significant difference in table 2 to 4, and removed one 0 in the subscript in table 5.